A Unified Model for User Identification on Multi-Touch Surfaces: A Survey and Meta-Analysis

AHMED KHARRUFA, Newcastle University

THOMAS PLOETZ, Georgia Institute of Technology

PATRICK OLIVIER, Newcastle University

User identification on interactive surfaces is a desirable feature that is not inherently supported by existing technologies. We have conducted an extensive survey of existing identification techniques, which led us to formulate a unified model for user identification. We start by introducing this model that (1) classifies existing user identification approaches in five categories according to the identification technology, (2) identifies eight characteristic identification system parameters, and (3) proposes a way for visualizing the system's characteristics as points on a radar chart to allow for quick comparison and contrast between systems. This model is then used to present our survey of existing user identification approaches and visualize their characteristics, highlighting their strengths and limitations. The model also makes it possible to visually represent requirements of systems that require user identification, identify existing approaches that can meet an application's requirements, and help report on and evaluate new approaches to user identification systematically.

CCS Concepts: • **Human-centered computing** → **Touch screens**; Collaborative interaction;

Additional Key Words and Phrases: Multi-touch Interaction, Interactive Surfaces, User Identification, Authentication, IMU, Wrist-worn Devices

ACM Reference format:

Ahmed Kharrufa, Thomas Ploetz, and Patrick Olivier. 2017. A Unified Model for User Identification on Multi-Touch Surfaces: A Survey and Meta-Analysis. *ACM Trans. Comput.-Hum. Interact.* 24, 6, Article 39 (December 2017), 39 pages.

<https://doi.org/10.1145/3144569>

1 INTRODUCTION

Direct touch-based interaction forms a central part of many ubiquitous computing applications (Bhowmik 2014). Recent reductions in the cost of multi-touch interaction sensing technology have seen the incorporation of multi-touch devices, in various shapes and sizes, into work, education, and public settings.

Multi-touch surfaces that are large enough to allow more than one user to interact with the surface simultaneously (e.g., tablets, tabletops, and multi-touch computer displays and public displays) naturally support collaboration. However, current mainstream and affordable sensing technologies

This work was funded by EPSRC awards EP/M023001/1.

Authors' addresses: A. Kharrufa and P. Olivier, Open Lab, Urban Sciences Building, Newcastle University, 1 Science Square, Science Central, Newcastle upon Tyne, NE4 5TG, UK; emails: {ahmed.kharrufa, Patrick.Olivier}@newcastle.ac.uk; T. Ploetz, Georgia Institute of Technology, School of Interactive Computing, Tech Square Research Building, 85 Fifth Street NW, Atlanta, GA 30308, USA; email: thomas.ploetz@gatech.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 1073-0516/2017/12-ART39 \$15.00

<https://doi.org/10.1145/3144569>

for interactive surfaces do not support the identification of, or even differentiation between, users interacting with the surface. In many multi-user applications, it is desirable to identify the user engaged in an interaction, or at least to distinguish between inputs from different users. For example, in educational contexts, having the ability to sense the level and type of contribution made by distinct students at a shared interface allows for the design of educational applications to explicitly promote collaboration, as well as online and offline analysis of interaction (Kharrufa Balaam et al. 2013; Kharrufa Martinez-Maldonado et al. 2013; Martínez et al. 2011; Martinez et al. 2011).

Other features made possible with user identification include the provision of different levels of access to actions based on role and privileges (Kharrufa et al. 2015), the provision of customized user interface experiences based on user preferences, supporting multi-user undo (Ryall et al. 2006; Holz and Knaust 2015), requiring consensus for performing critical commands or for coordinating collaboration (Kharrufa et al. 2010), detecting and supporting handedness (Zhang et al. 2012), and as a form of authentication (Holz and Knaust 2015).

Accordingly, a number of different approaches to user identification at multi-touch interfaces have been proposed (e.g., (Dietz and Leigh 2001; Holz and Knaust 2015; Ramakers et al. 2012; Roth et al. 2010)). However, existing identification techniques have one or more of the following practical limitations: (1) dependence on specific touch-sensing technologies (e.g., (Dietz and Leigh 2001; Holz and Knaust 2015; Roth et al. 2010)); (2) the use of impractical infrastructures (e.g., (Dohse et al. 2008; Ramakers et al. 2012; Walther-Franks et al. 2008)); (3) dependence on specific certain screen sizes; (4) restrictions on the mobility of users while interacting with a surface (e.g., (Dietz and Leigh 2001; Dohse et al. 2008)); (5) restrictions on the positions of interaction on the surface (e.g., (Scott et al. 2004)); (6) or differentiation rather than identification of users (e.g., (Annett et al. 2011)).

For a practitioner wanting to design or choose a system that supports user identification, making a decision between existing systems is not easy. There is no agreed upon criteria to compare and contrast between these systems or even an agreed upon terminology for describing their characteristics. What adds to the challenge is that the descriptions of many of these systems focus on how well they perform in the specific contexts they are designed for. This makes their adaptability to, and limitations within, different settings rather vague; especially for non-technical practitioners. This work is motivated by our experience in working on a number of technology interventions that required user identification on multi-touch surfaces and the lack of a common scheme to examine existing systems.

Accordingly, our objective and contribution in this work is to (1) develop a unified conceptual model for direct-touch, multi-user identification systems, which identifies the common parameters that characterize such systems and allows the visualization of the values of these parameters in a way that highlights the systems' strengths and limitations; (2) use the model to classify existing user identification systems; and (3) present an extensive survey of existing approaches for user identification.

The development of the unified model and the review of existing work were done as an iterative process. We started with the literature review to inform the first version of the model that was then validated and refined through re-examining the reviewed systems. We first start by introducing the model (independently from existing systems) and then use the model to structure and guide the review of related work.

2 OPPORTUNITIES AND APPLICATIONS OF USER IDENTIFICATION

The ability to identify, or at least distinguish between, users interacting with multi-touch surfaces opens up a new design space for multi-user multi-touch applications that have implications

ranging from the nature of graphical user interface elements and gestures, to higher level features of applications.

2.1 User-aware GUI Elements

Ryall et al. (2005, 2006) presented a conceptual framework that articulated the benefits and opportunities for customization that arise when Graphic User Interface (GUI) elements are supplemented with user identity (iDwidgets). This framework had four dimensions, function, group input, content, and appearance as follows:

- Customizing *function*: User interface elements behave differently depending on the identity of the user interacting with them. Examples include buttons that perform different functions depending on who pressed them (e.g., an undo button undoing actions only of the person who pressed it); semantic interpretation (e.g., typing “dad” in a photo search box returns different results depending on who performed the search); differentiated behavior (e.g., a paint brush varies its brush thickness depending on the prescribed user’s preferences); and privileged access (e.g., widgets responding only to specific users, or subset of users, depending on the predefined privileges).
- Customizing *group input*: Individual interface elements can support cumulative effect (e.g., a command button that does not trigger unless a number of distinct users sequentially press it); simultaneous input (e.g., a critical command button that only triggers when a number of privileged users press it simultaneously); modal input sequence (e.g., upon selecting an object, the system reacts differently based on the operation mode of the user who selected it); and audit and logging (e.g., for data analytics).
- Customizing *content*: User-specific lists, such as bookmarks and recently used commands, show different content depending on the identity of the user that activates them.
- Customizing *appearance*: Properties (e.g., language) and aesthetics (e.g., colors) can be customized to different users.

2.2 User-aware Gesture Event Handlers

Users’ interactions are not limited to GUI elements, and in many cases users interact directly with digital content such as documents and pictures. Morris et al. (Morris et al. 2006a; Morris et al. 2004; Morris et al. 2006b) explored how user identification makes it possible to move away from relying on social protocols to resolve conflicts arising from multiple people working on a shared surface, to resolving them through supporting cooperative gestures (“interactions where the system interprets the gesture of more than one user as contributing to a single, combined command” p. 1201 (Morris et al. 2006b)). The benefits of such cooperative gestures mirror those of the *function* and *group input* dimensions of user-aware GUI elements. However, interacting directly with digital content can sometimes afford more natural workflows. Holz and Knaust (2015) provide a compelling example of temporal access control for a digital object in their biometric touch sensing surface, for which a user can give temporary access to her document by simultaneously touching it with the person requesting access.

2.3 Example Applications

User identity at the GUI element and gestures levels has the potential to support the design of richer multi-user applications for interactive surfaces. For example, Digital Mysteries (Kharrufa et al. 2010), a collaborative learning application built for an interactive tabletop, supports user identification through the use of pen-based interaction. The ability to identify users and their interaction was exploited in a number of application features that promoted collaboration, such as

requiring all students to confirm approval at different critical points in the process (like creating and naming groups or progressing through stages), and visualizing levels of participation. Furthermore, since all logged actions carried user identity information, it was possible to analyze the different sequential activity patterns for different groups to identify the difference in these patterns between higher achieving and lower achieving groups (Martinez et al. 2011) as well as to have interaction visualizations of these groups at the user-level (Al-Qaraghuli et al. 2011).

In the social domain, Piper et al. (2006) developed SIDES (Shared Interfaces to Develop Effective Social Skills) using DiamondTouch (Dietz and Leigh 2001) to support adolescents with Asperger's syndrome to practice effective group work skills using a four-player cooperative computer game. User identification allowed the exploration of the effects of human- vs. computer-enforced rules and the embedding of user-aware control structures into the design.

For multi-user applications requiring reliable enforcement of different levels of privilege, Holz and Baudisch (2013) demonstrate the benefits of their fingerprint based user identification system in supporting authorization of actions such as approving invoices and associating payment limits to different users depending on their level of authority. For gaming, a typical application of user identification support is to prevent players from taking actions on behalf of the other players (Bianchi and Je 2017).

Schmidt et al. (2010c) gave a number of example sub-tasks that benefit from user identification using their IdLenses interaction technique (which builds on HandsDown (Schmidt et al. 2010b) approach for user identification). Their examples include automatic filling of user credentials when hovering the IdLense over the login details area, supporting user-level rather than application level clipboards, and supporting personalized annotations on shared documents.

3 A CONCEPTUAL MODEL FOR USER IDENTIFICATION

It is useful at this stage to start with defining some of the key terms related to user identification as used in this article.

- *User identification* is used when the system can recognize the users interacting with it. This can range from simply linking all interactions during session to an ad hoc identity entered by the user for that session, to securely, persistently, and reliably identifying a user identity based on a predefined database. Systems that rely on unique biometric information such as finger prints fall into this category (e.g., (Holz and Baudisch 2013; Schmidt et al. 2010b)).
- *User differentiation* (or *distinguishing between users*) is when the system can only recognize that different input gestures have been carried out by different users, regardless of who these users are (i.e., the system cannot link touches to specific users). Systems that rely on the position of interaction or the position of the users to differentiate between touches fall into this category (e.g., (Scott et al. 2004; Annett et al. 2011)).
- *User authentication*, while not the focus of this work, refers to when the system requires the user to confirm a claim of identity. Entering a password to confirm identity after entering a username is a typical example of user authentication. Authentication only requires comparison of some input against a specific template (that of the user who made the claim), while user identification requires comparing input to a database of users.

For the remainder of this article, we use the term *user identification* where we are not seeking to necessarily distinguish among identification, differentiation, or authentication.

Most of the existing approaches for user identification (or differentiation or authentication) for multi-user touch surfaces have been developed either for specific hardware, to solve a specific problem, or to respond to a specific set of requirements. As such, it is not always easy for

practitioners and system designers who require some level of user identification to decide on the option most suitable for their requirements. There is a need for a conceptual model that

- (1) Allows for systematic categorization of existing work on user identification on multi-touch surfaces;
- (2) Identifies the common parameters that help characterize and describe a user identification approach (or express a system's requirements);
- (3) Identifies the range, or discrete set, of values associated with each of the identified parameters;
- (4) Allows for visualizing the values of these parameters for quick visual comparison between different methods, or for identifying approaches that best satisfy specific system requirements.

We have conducted an extensive review of related work in the past two decades to inform a unified conceptual model that addresses the above four points. Important to developing the model was the categorization of existing systems according to the defining features of the approaches used to achieve user identification. A number of themes were identified, which were then reviewed by the research team and grouped into five main categories that cover all the following reviewed systems: (1) sensing technology inherent; (2) overhead camera-based tracking; (3) zone-based; (4) wearables and handheld; and (5) biometrics. Each category has advantages and limitations that are characteristics of the approach used and are independent from the specifics of the different implementations. Starting by identifying the *category* of a user identification system gives an indication of its general strengths and limitations before going into the specific details.

Across categories, eight *characteristic parameters* that apply to all systems can be identified as follows: (1) infrastructure; (2) touch sensing dependence; (3) user differentiation; (4) points of interaction; (5) instantaneity; (6) user mobility; (7) interaction position; and (8) surface size. The identification of these characteristic parameters, their values, and the way to visualize them was an iterative, inductive process. It involved developing a version of the model, then refining it through testing it against reviewed systems and discussing it with a research focus group of 5–7 researchers with an experience in user identification and interactive surfaces. The process went through four major iterations.

Iteration 1 (Figure 1(a)): The first version was a simple three-dimensional model that consisted of the typical time and space parameters in the X–Y horizontal plane as well as error tolerance parameter as a Z-axis. The time parameter was divided according to when identification occurred with respect to user actions: a pre-action identification process (for techniques that require some kind of identification before taking an action such as HandsDown (Schmidt et al. 2010b)), during action, post action, or offline (after the completion of a session). The space parameter was divided into identification through interacting in personal spaces, across a shared space, or using multiple-physical spaces. The application of the model to a number of identification approaches such as DiamondTouch, camera-based, and wearable-based approaches during the research group discussion revealed limitations in reflecting some key system characteristics (e.g., level of dependency on touch sensing hardware), difficulty in understanding the time parameter as well as difficulty in visualizing systems using this three-dimensional model.

Iteration 2 (Figure 1(b)): The main decisions for the second version are as follows: (1) The use of a visualization that is more flexible in terms of number of axis that can be used, that does not require a unified scale for all its axes, that can create signature visual representation for the different systems, and that supports visual comparison between different systems or between systems and requirements. A radar chart representation was selected as it meets all these requirements. To allow for comparing between systems (Razmi et al. 2000), the points on the axes should be selected so that

Fig. 1. First three iterations of the conceptual model and its visualization. (a) A three-dimensional model for version 1. (b) A 5-parameter model represented as a radar chart. (c) A 6-parameter model (that excluded the accuracy parameter) represented as a radar chart.

moving away from the center of the chart reflects systems that are more flexible or more effective at user identification. If a system and a set of requirements are visualized on the same graph, then the system meets the requirements if its plot encompasses the plot of the requirements. (2) Include an additional hardware independence parameter and split the time parameter into two. This model was again tested against a wider range of systems. A research group meeting concluded that the radar chart representation was suitable to visualize, and allow the comparison between, systems and that there is a need to further refine the parameters as they were still limited in representing all the main characteristics of the identified user identification systems.

Iteration 3 (Figure 1(c)): One of the key decisions and discussion points at this stage was the removal of the accuracy parameter. The main reasons are that the level of accuracy is mainly determined by the system implementation, the hardware used, and the way accuracy is defined and measured and not by the specific characteristics of the approach. Moreover, for many systems, no measure of accuracy was reported, and unlike the other parameters, accuracy of a system is not something that can be figured out from an understanding of how a system works. Accordingly, it is not possible to use accuracy as a true representational feature of a system (independently from

Fig. 2. Examples of the five main categories for user identification based on the technology or technique used.

a specific implementation) and as a comparison metric between systems. Another conclusion was the need for parameters that represent infrastructure (to distinguish between systems like DiamondTouch, camera-based systems and wearable systems for example), user differentiation (to be able to distinguish between systems that only support user differentiation from ones that support user identification), and points of interaction (to distinguish between the levels of support for user identification from a single finger to multi-finger two handed interactions), and the need to change the two time related parameters into instantaneity axes and spatial/temporal constraints parameters. This model was again tested against the growing set of identified literature. This exercise highlighted that the temporal constraints are reflected in the instantaneity axes and that the space constraints were too general and can be specified further in terms of user mobility, position of interaction, and the supported surface size. This lead to the development of the fourth and final version that made it possible to describe, compare, and visualize all the identified literature. The iterative process also looked at and refined the key values associated with each of the identified parameters. The unified conceptual model is used to both structure the survey of related work that follows and provide a common terminology for describing that work. The model will thus be presented first in abstract terms without reference to existing work.

3.1 Categories of User Identification Approaches

We have identified five categories that define the main features of multi-touch user identification systems (Figure 2).

- *Sensing technology inherent*: The identification mechanism is built into the sensing technology not requiring any external wearables or sensors in the environment (Figure 2(a)).

- *Overhead camera-based tracking*: Computer vision-based systems that rely on overhead cameras for body and hand tracking and consequently user identification (Figure 2(b)).
- *Zone-based*: Systems that achieve user identification based on users interacting within specific zones on the surface (Figure 2(c)) and/or on the users' positions around the surface.
- *Wearables and handheld*: Systems that achieve user identification using wearables (e.g., rings, wristbands (Figure 2(c)), or gloves) or handheld devices (e.g., mobile phones).
- *Biometrics*: Systems that rely on user biometrics (e.g., finger-print (Figure 2(d)) or impedance profile) for user identification.

These categories cover all devices and user identification approaches described in this work and they should be general enough to cover all future systems as well. The devices include all surfaces that are *large enough* to allow for more than one user to interact with them simultaneously (starting from a tablet size and excluding smart phones). This also means that the devices must support *direct, multi-point simultaneous* interaction with their surfaces (ideally through direct touch but interaction through other means such as digital pens is possible). This includes existing commercial tablets and multi-touch computer displays (which normally use capacitive sensing technology), the different commercial and research tabletop implementations (which vary greatly in the sensing technology used), and large interactive wall displays (which again vary in the sensing technology used with some supporting special digital pens only, some supporting touch only, and others supporting both).

3.2 Characteristic Parameters

Through the iterative process of developing the model, we have identified eight characteristic parameters that apply to all analyzed systems that can be used to describe their main advantages and limitations. During the process and through examining existing literature, we also identified the set of ordinal values possible for each parameter and a number of optional qualifiers providing further details for some of these values. The table for each characteristic below states its key possible values. If any of these values require a qualifier to provide further information, these are presented in parenthesis. For example, when stating that a certain user identification approach is technology dependent, the qualifier in parenthesis describes the type of technology it is dependent on (capacitive, infrared vision-based, etc.).

- *Infrastructure*: This refers to the placement of technology used for user identification. Ideally, identification related hardware should be embedded in the sensing technology thereby removing the need for any external hardware to support identification (such as an overhead camera). In reality, this is not always the case and existing systems use one or a combination of the techniques as below. For the infrastructure parameter, a qualifier is needed to further state the specific type of technology used (as in the examples below).

Environment	Worn/handheld	Embedded
External to the surface (overhead cameras, proximity sensors).	Users wear a device (ring, gloves, wristband) or hold a device (a digital pen or a mobile phone).	The identification technology is part of the touch sensing technology (fingerprint aware screen).

Infrastructures that rely on external hardware are generally the most limiting as they are the least portable and may constrain the settings in which they can function properly (e.g., lighting conditions and space requirements). Worn/handheld allow more flexibility, but they still require users to wear/hold additional hardware. Embedding the identification in the sensing technology itself overcomes these limitations.

- *Touch sensing dependence*: Identification should be possible independently of the touch sensing technology used. However, existing systems fall into one of the following three classes:

Custom	Dependent	Independent
Works only with custom developed surfaces, thus the most restrictive.	Works only with certain “existing” technologies (capacitive, vision-based). However, a system that works with widely used sensing technologies like capacitive screens is preferred over less common ones such as vision-based systems.	Works irrespective of the touch sensing technology used.

- *User differentiation*: Systems vary in their support of user identification ranging from uniquely and reliably identifying a person across devices, to only having the ability to differentiate between multiple touches on the same surface. Systems that claim to have user identification reliable enough for usage in secure or mission critical applications will be further identified with the (+) qualifier.

An individual (differentiation)	Unique individual(+) (identification)	Unique across devices(+) (identification across devices)
The system differentiates between users but does not identify them.	Users are uniquely identified on one device but not when moving between devices.	Users are uniquely identified even when switching between devices.

- *Points of interaction*: Users should be able to use two-handed, multi-finger interaction and still be identified. Ideally the system should even distinguish between left and right hands, as well as distinguish between all fingers. However, some systems only allow one-handed, single-finger identification. A system can be within a range with the following key points:

One-handed, one finger	One-handed, all-fingers(+)	Two-handed(+), one-finger	Two-handed(+), all-fingers(+)
	The (+) qualifier means it is possible to distinguish between fingers.	The (+) qualifier means handedness support.	

- *Instantaneity*: Identification should occur instantly upon touching the surface. In reality, most systems require some complicated processing to achieve identification. We use the term online identification to describe a system that still gives the feel of real-time identification, and offline for systems that can only identify users after the completion of the interaction. A system can be within a range with the following key points:

Offline	Online	Instant
Works on offline data after the completion of an activity rather than in real time.	User identification occurs within a short delay after touching the surface that is still usable for scenarios requiring real-time user identification	Instant user identification upon touching a surface.

The final three parameters relate to the spatial constraints of the system.

- *User mobility*: Users should be able to move freely around the interactive surface and still be identified. However, some systems require users to either stay in a certain position, or even ensure contact with a chair or a floor mat during the interaction. Consequently, systems fall on any point in a scale whose two end points are as follows:

Constrained	Free
Users need to maintain a certain position, and sometimes posture (e.g., seated) around the surface.	Users can move freely around the surface.

- *Interaction position*: Users should be able to interact with surface at any position and still be recognized. However, some systems rely on associating certain zones on the surface with different users to achieve user identification. The main points on this scale are as follows:

Fixed	Constrained	Anywhere
To achieve identification, users must interact with the surface only within certain zones that have been pre-defined and linked to users.	Users can dynamically control the zones that are linked to them. The flexibility and ease of controlling these territories vary based on the approach used.	Users can interact anywhere on the surface and still be identified.

- *Surface size*: While ideally, a user identification technique should work regardless of the surface size, some impose restrictions of the size of surface supported. For example, those that rely on overhead camera for body tracking require a large surface (large as in interactive wall displays or tabletops). On the other hand, those that rely on expensive dedicated hardware may be limited to small surfaces (as in tablet size). The range for this parameter is thus given as follows:

Restricted (large or small)	Any
Restrictions on supported surface size. A qualifier is needed to specify the type of size restriction. For example (large: tabletop) or (small: tablet).	No restrictions on surface size.

Fig. 3. 8-axis radar chart conceptual model for user identification on multi-user, multi-touch devices. The labels show the possible ordinal values for each parameter.

3.3 Visualization

Our conceptual model makes it possible to visually express the main characteristics of user identification techniques (as extracted from the literature) and the requirements of a specific system as plots of an 8-axis radar chart (Figure 3). The ordinal values are mapped to these axes such that the most flexible (or desirable) values are placed further away from the center of the radar chart and the most restrictive are placed near the center of the chart (Razmi et al. 2000). It is important to point out that as we are only using ordinal values, when comparing systems, the exact position on the axes is not as important as the relative position thus these values are simply distributed evenly along the axes. When comparing systems, placement of points for “online” (instantaneity axes), “size restriction” (surface size axis) and “constrained” (position and mobility axes) should reflect how fast (in case of online), restricted (for surface size) and constrained (for mobility and position) a system is compared to others. For some values a qualifier is necessary, which needs to be clearly stated on the chart in parenthesis (e.g., *capacitive* for hardware dependence, and *small/large* for surface size). Following these rules, when visualizing a certain system’s requirements in terms of user identification, any technique whose plot entirely contains that of the desired system, and whose qualifiers match or are more flexible (as in type of hardware dependence or size limitations),

is considered a viable option. This also makes it easy to compare different systems on the chart as it quickly illustrates the different systems strengths and limitations.

The axes on the radar chart are placed so that related ones are placed next to each other. *User mobility* around the interactive surface, *interaction position* and *surface size* axes are “space” related and are placed next to each other. *Infrastructure* and *touch sensing dependence* are “hardware” related and are placed next to each other. *User differentiation* and *points of interaction* are related to the level of “precision” of user identification and are placed next to each other. Moreover, the *infrastructure* used normally influences the supported *surface sizes*. For example, systems that have an infrastructure embedded in the environments (e.g., camera-based or proximity-based systems) cannot support small surface sizes such as tablets while systems that use some sort of wearable technology can normally support any surface size. For that reason, *infrastructure* is placed next to *surface size*. Likewise, most systems that rely on embedded *infrastructure* use a custom built *touch sensing* hardware (another reason for placing *infrastructure* and *touch sensing dependence* axes next to each other). Apart from *surface size*, *infrastructure*, and *touch sensing dependence*, there generally is no correlation between the other parameters.

4 A SURVEY OF USER IDENTIFICATION WORK

We have conducted an extensive and iterative review of previous systems that have tackled the problem of user identification on multi-user/multi-touch surfaces over the past two decades. The process for identifying relevant work to include in this survey can basically be divided into three phases:

Phase 1: Identifying a core set of papers. We searched the ACM digital library for all papers published at the relevant venues (CHI, UIST, ITS, and TEI) since 1997. The search used the following keywords “user identification,” “identify users,” “user differentiation,” “differentiate users,” “touch identification,” “touch differentiation,” “differentiate touches,” “finger identification,” “distinguishing user,” “distinguishing between users,” “user authentication,” “multi-touch authentication,” “multitouch authentication.” We used the default “Any field” search during the specified time filter. This returned 734 results. A first pass through the titles and abstracts of these papers reduced the list to only 55 papers. A second more in-depth pass through the remaining 55 papers by a specialized member of the research team reduced the list to only 17 highly relevant papers. The following criteria were applied to all papers to decide on which ones to exclude from the list:

- Exclude any work that does not target *direct touch interaction* or that could not reasonably be applied to direct touch interaction.
- Exclude any work that does not contribute a technology or an approach to support some sort of *user identification or differentiation* (or that could not reasonably be applied to user identification/differentiation) in a *simultaneous multi-user interaction* context.
- Exclude any work that cannot directly or indirectly support *continuous identification/differentiation* of users.
- Exclude any work that applies only to small mobile devices (i.e., smart phones) and that cannot extend to support multi-user interaction around *large devices* (tablets and larger)
- Exclude any work that is specific to secure user authentication and that cannot be easily extended to support continuous simultaneous multi-user interaction.
- Exclude any work that is aimed at the more general problem of tracking users in 3D space and that does not specifically address user identification around touch devices.

This phase was not expected to produce a comprehensive list but to identify a core set of papers to work from. We added a number of relevant papers that the research team were already aware of to this core set.

Phase 2: Cross referencing. During this phase, we looked at all the papers cited by the core set of papers as well as the papers citing the core set of papers. The same exclusion criteria as in phase 1 were used.

Phase 3: We continuously repeated the process in phase 2 to the newly identified papers until no new relevant papers are identified.

We used the review process to both inform and validate the different iterations of the conceptual model presented in terms of main categories, characteristic parameters, and possible values for each of these parameters reaching to the final design. We then validated the final model by using it to re-examine the literature. The model was used to provide structure, common terminology, and a visual representation of the reviewed literature.

Before presenting the survey of systems that address the challenge of user identification, it is useful to briefly discuss the only attempt that the survey has identified to categorize user differentiation techniques on interactive surfaces. Sahdev et al. (2017) introduced a simple three dimensional framework that aim to help categorizing techniques that support user differentiation. The three parameters they identified are *precision* (only supporting basic *tracking* or identification/differentiation on the *user-level*, *hand-level*, and *finger-level*), *persistence* (working at the *input frame level*, *touch point level*, *touch grouping level*, *session level*, and *lifetime*), and *mobility* (supporting identification/differentiation at the *room*, *tabletop*, *tablet*, *phone*, and *watch level*). This is represented as a two-dimensional table with the third mobility parameter visualized using color codes inside the two-dimensional table. Examining these parameters in relation to our model and excluding the values that fall outside the range that allow user identification (supporting tracking only, or working at the input frame or touch point level) show that the points in the *precision* and *persistence* parameter are represented in slightly different terms in our model using the *points of interaction* and the *user differentiation* parameters. The mobility parameter maps almost directly to the size parameter but as we are focusing specifically on “interactive surfaces” that allow “multi-user” interaction, we have excluded work that looks at the watch, phone, and room level. While this basic framework is not as general as our proposed model due to its limitation in identifying important features such as hardware independence, type of infrastructure used, and the limitations imposed on position of interaction and users’ mobility, it does emphasize the need for models to help in categorizing, comparing, and visualizing different user identification systems.

To illustrate the value of our model, with its visualization, we present two possible scenarios that require user identification then use the model to show the level to which the reviewed systems meet the requirements of these scenarios.

Scenario 1: Tablet-based classroom applications are growing rapidly as reported by British Educational Suppliers Association (BESA) (2013) and are seen by many as one of 21st century’s key classroom innovations. Accordingly, a number of classroom-oriented collaborative learning tools have been developed for tablets that support two simultaneous users working around the tablet (e.g., (Rick and Jochen 2012; Thinking n.d.)). Such collaborative learning tools can benefit from user identification not only to support accountability but also to show level of participation and support collective agreement on different decisions in the process for example (Kharrufa et al. 2010). For this scenario, we can imagine pairs of students collaborating in a classroom around off-the-shelf tablets (i.e., a “small” surface that depends on “capacitive” sensing technology); it is only required to differentiate between pairs of students on each tablet for accountability of collaboration (knowing who is doing what). Students will be seated, so their mobility is constrained. Only one handed, one finger interaction is expected. Some commands require two handed gestures, but it is not critical to identify which student performed such gestures. These requirements can be visualized using the conceptual model as in Figure 4(a). The visualization quickly shows the minimum

Fig. 4. (a) Minimum requirements for Scenario 1 – classroom-tablets. (b) Minimum requirements for Scenario 2 – classroom-interactive display.

requirements for any user identification system to be useful for this scenario as follows: must be able to support small (tablet) surface size; infrastructure can be embedded or worn (if necessary), must support capacitive touch sensing, it is enough to differentiate between users using one-hand, one-finger for interaction; identification needs to at least be online; user mobility can be restricted as students will be seated and, while in most cases identification should be supported anywhere

Fig. 5. A typical DiamondTouch setup: front-projection onto a tabletop surface. Users are seated on chairs that are connected to the tabletop surface (Esenther et al. 2002).

on the surface, zone-based identification may be acceptable for some applications (e.g., ones that split the surface into two halves, one for each student).

Scenario 2: Interactive whiteboards are becoming increasingly popular in modern classrooms. They are seen as one of the technologies that can greatly enhance teaching and learning (Heemskerk et al. 2014). Modern interactive displays are moving away from the original single pen-based interactive boards to ones that support multi-pen and multi-touch interaction. For this scenario, we can imagine a classroom with a large interactive display on the wall that needs to distinguish between teacher's input and students' input to allow the teacher privileged access to some commands. If more than one student is interacting with the display at the same time, the display needs to differentiate between inputs from different students. It will be especially useful if the display can uniquely identify each student. These requirements are visualized using the conceptual model in Figure 4(b) which again shows the minimum requirements needed as follows: must support large (wall display) surface size; environment-based identification infrastructure is acceptable; custom touch sensing technology is acceptable; must differentiate between student and identify the teacher as a unique individual; one handed, one-finger identification is acceptable; identification needs to at least be online with no restrictions on user mobility or interaction location.

The review that follows is structured based on the five *categories* of user identification approaches. The conceptual model is used to inform the review of the different systems, and then the radar chart visualization is used to plot the key features of each system, or set of related systems. The visualization of the two scenarios in Figure 4 is shown next to these plots for reference. A “checkpoint” subsection is then used to demonstrate how the visualizations can be used to examine the extent to which these systems meet the two scenarios' requirements.¹

4.1 Sensing Technology Inherent

DiamondTouch (Dietz and Leigh 2001; Esenther et al. 2002) uses a front-projected surface with an array of RF-emitting antennas to support user identification (Figures 5 and 6). Touches are

¹Visualizations of all the systems covered in this survey can be downloaded as SVG vector graphics from <https://github.com/ankharrufa/umodel-tuid.git>.

Fig. 6. Visual representation of user identification features of DiamondTouch. The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. The plots show that DiamondTouch does not meet the requirements of either scenario.

capacitively coupled through the user to a receiver connected to the user's chair that allows the user initiating the touch to be identified instantly and reliably. The receiver can also be conductive floor plates that the user can stand on. DiamondTouch, thus, uses *embedded* technology that require rather costly *custom* hardware. It only supports *user differentiation* rather than identifying them as unique individuals, but allows for *instant, two-handed multi-finger (cannot distinguish hands and fingers)* interaction. Due to the need to maintain contact with the chair, or floor plate, DiamondTouch significantly *constrains users' movement*, but allows *interaction anywhere* on its surface. DiamondTouch is only available as a *large surface*.

Checkpoint: Figure 6 shows the key features of DiamondTouch. The blue dashed plot and the red dotted plot show the plots for scenarios 1 and 2 for comparison. The plots show that DiamondTouch is not suitable for this scenario as it does not work on standard capacitive displays but require a custom-built hardware. Another limiting factor is that DiamondTouch is only available in large (tabletop) form factor in contrast to the tablet size surface requirement. For Scenario 2, the plots show that while DiamondTouch's large size makes it possible to use it as a wall display,

it does not meet the requirements of uniquely identifying users (identify the teacher) and that of free mobility.

Sahdev et al. (2017) presented GhostID that can differentiate in a non-persistent way between whether multiple simultaneous touches are by different fingers of the same hand, the other hand of the same user, or by different users. This means that the system cannot differentiate between separate non-overlapping touches in a session. While this is a significant limitation making GhostID not suitable for most practical applications that require user identification (including our two educational scenarios), the system still allows for a set of interaction techniques that can take advantage of the system. GhostID uses an *embedded* infrastructure based on a *custom capacitive sensor* that, other than the non-persistence limitation, is free from *size*, *position*, and *mobility* constraints.

4.2 Overhead Camera-based Hand Tracking

To overcome the limitation of using custom hardware and support user identification independently of the sensing technology, a number of researchers have explored the use of overhead camera systems (Dohse et al. 2008; Clayphan et al. 2013; Martínez et al. 2011; Murugappan et al. 2012; Ramakers et al. 2012). All these systems use *environment-based infrastructure* (the overhead camera), support identification that is *independent from the sensing technology* used, support *on-line* identification, and allow *interaction anywhere* on the surface but are limited to *large surfaces*. However, they differ in terms of user differentiation, points of interaction, and their constraints on user mobility.

Dohse et al. (2008) use an overhead camera and computer vision to track users based on skin color segmentation techniques to *differentiate between users* around tabletops (but not as unique individuals) (Figure 7(a)). They use an inexpensive 640×480 web cam with a frame rate of 16 fps for hand tracking on top of a large FTIR-based tabletop. The system *supports two-handed, multi-finger* interactions. The use of the overhead camera also allows detection of interactions above as well as on the display. However, the system has several limitations: (1) It can only associate touches with users as long as only one user stands at each of the four edges of the table thus *restricting users' mobility*. (2) As with most vision-based systems, users must avoid overlapping hands to prevent interfering with the tracking. (3) A limitation very specific to this system is that, to facilitate accurate tracking of hands, applications need to maintain a reasonably constant color intensity throughout the display. No formal user study was conducted that provide quantitative evidence for the identification accuracy of the system.

Murugappan et al. (2012), Clayphan et al. (2013) and Martínez et al. (2011) use depth images generated by an overhead depth camera (Microsoft Kinect) to *differentiate* between users. Both systems *support two-handed, multi-finger* interaction. Murugappan et al. (2012) report very briefly on the ability of the system to differentiate between users using heuristics to model users' positions around the table based on the partially visible users' hands (Figure 7(b)). This means that users need to maintain *fixed relative positions* around the table and not be too close to each other. However, their system supports *handedness* (which hand) and identifies common postures, and thus associates touches with specific fingers for known postures.

To overcome such limitations, Clayphan et al. (2013) and Martínez et al. (2011) rely on tracking the position of each user's body (mainly the user's head) and use an algorithm to trace a line between any touch on the surface and the head of the user responsible for that touch (Figure 7(c)). This allows for recognizing *two-handed* (*not clear if handedness is supported as well*), *multi-touch* interactions and even interactions using pens. This also means that users can *move around the table but with limitations* (as long as their heads are within the tracking region of the depth sensor). The authors report on extensive performance testing for the system in Clayphan et al. (2013). Accuracy of identification varies widely depending on the testing condition. 98–100% accuracy is achieved

Fig. 7. Overhead camera-based hand tracking systems. (a) Using skin color segmentation technique (Dohse et al. 2008). (b) Using heuristics to models users’ positions around the table based on the partially visible users’ hands (Murugappan et al. 2012). (c) Using body position tracking with an algorithm to link touches on the surface to corresponding users’ heads. (d) Detecting unique features from the back of users’ hands as identifiers using a high-resolution camera.

for two to six users interacting close to their tracked body position, but when touches are near other users, accuracy drops to 81–97%. Arm crossing also negatively affects identification with accuracy falling to the range of 75–92%. A seated testing condition is reported to have 3% better accuracy than standing testing condition. The two main causes of errors are as follows: (1) a user other than the one who performed a touch was closer to that touch, and (2) touches by different users are in close proximity to each other. That is, the system has *limitations on the position of interaction*.

While still using an overhead camera for tracking, Carpus (Ramakers et al. 2012), allows for more flexible user identification by leveraging the back of users’ hands (dorsal region) as

identifiers (Figure 7(d)). By using a high-resolution camera, the system can capture unique identifying features from a hand's visible region during an interaction. These identifying features are compared to a database of user-feature-pairs that is constructed beforehand during training sessions that allows for *limited unique identification of pre-registered users*. The system is *not designed to offer authentication for privacy or security*, but registration is purely needed to support identification. While the technique can, in theory, uniquely identify a user across devices, the infrastructure used (overhead cameras and a large surface) makes this impractical thus limiting it to a single tabletop in practice.

This approach enables users to *freely move around the table* during the session, and can differentiate between actions performed by a user's right hand and left hand. It can also identify touches if the arms overlap as long as the back of the hand remains detectable by the overhead camera. However, when performing very fast movements, the camera may not capture sharp enough images to allow for identification which puts *some constraints on its online identification performance*. A user study showed that the system uniquely identified both hands of users with more than 97% accuracy for up to 20 registered users. However, accuracy drops to 82% (as a worst case scenario) for complex hand postures that do not allow for clear capture of the dorsal hand region. Accordingly, the system allows for *two-handed+, multi-finger identification (with some limitation on the multi-finger gestures)*.

Checkpoint: Figure 8 shows the range within which overhead camera-based hand tracking systems operate. The plots show that these systems fail to meet the needs of Scenario 1 (blue dashed plot) as they require external hardware setup (overhead cameras) that is impractical for a classroom and only work with large surfaces. On the other hand, the plots show that systems that support identification of users as unique individuals and support user mobility completely contain the plot for Scenario 2 (red dotted plot), thus can meet the desired requirements. If all students are pre-registered, then all students can also be identified uniquely.

4.3 Zone-based

The simplest and most basic approach to identify users on a large interactive surface is to make use of the zonal (or territorial) division of the interactive space (Scott et al. 2004). When working around a large interactive space, and more so around tabletops, the space is implicitly divided to include a personal space for each user, a public (or group) space in the middle, and less well defined (in terms of position and ownership) storage territories. If these implicit spaces are made explicit, and assuming users do not interact within the personal spaces of others, then it is possible to associate all interactions starting from within a user's personal space to that user. The main limitation of such a zone-based approach is to *interact only within a fixed region* (which accordingly enforces limitations on application design and, depending on number of users, display size). Moreover, it only *differentiates between users* rather than uniquely identifies them. However, this approach can be considered as *embedded* as it does not require any external infrastructure, is *independent of the sensing technology*, allows *2-handed, multi-finger interaction (but obviously without distinguishing between hands and fingers)*, *works instantly*, and *supports users' mobility*.

Checkpoint: Figure 9 shows that zone-based identification only covers the requirements for Scenario 1 for applications that are designed so that students interact in their own specific regions. For Scenario 2, Figure 9 shows that zone-based systems fail to meet the requirements of interacting anywhere on the surface and uniquely identifying the teacher.

Medusa (Annett et al. 2011) is a proximity aware system that uses 138 IR proximity sensors distributed in three rings around a multi-touch tabletop (Figure 10(a)). The system works *independently of the touch sensing technology* as its identification infrastructure is in the *environment*

Fig. 8. User identification for overhead camera-based systems extends between the two gray regions depending on the specifics of the approach used. The blue dashed plot represents the minimum requirements of Scenario 1 while the red dotted plot represents the minimum requirements for Scenario 2. As overhead camera systems use an environment-based infrastructure and only support large surfaces, they cannot support Scenario 1. However, the red dotted plot for Scenario 2 falls entirely within the region covered by systems that support user mobility and unique identification of users.

(a frame around the table) rather than embedded or worn. The system associates touches with specific users and hands based on computing an “arm projection” area using the user’s identified body position and orientation and tracking of arms above the table (Figure 10(a)). When touches fall into overlapping “arm projection” area, the system relies on the minimum distance that exists between the touch point and each of the arm projections. Accordingly, medusa uses a zone-based approach, where the zone is dynamic based on the arm projection area (*constraints on interaction position*). By detecting the position of a user’s arms, it is possible to distinguish between right and left arms. Medusa thus supports *online*, *two-handed+*, *multi-finger* interaction with *user differentiation*, rather than unique identification, but *restricts users’ mobility* around the table. The design of the system also limits this approach to *large surfaces*. The reported limitations of this approach include cases where users cross paths as the body tracking algorithm may mislabel users when they separate, another limitation is apparent when users’ hands hover over the same area. The

Fig. 9. Zone-based identification allows interaction only in pre-defined personal spaces. The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. The plot for Scenario 1 falls within that of zone-based identification for applications that restrict interactions to personal zones. The plot for Scenario 2, however, does not fall within region covered by zone-based identification systems.

Fig. 10. (a) Medusa, a proximity aware system implemented on a tabletop (Annett et al. 2011). (b) Zhang et al. system uses the orientation of a touching finger, obtained from the touching hand's contour, to associate a touch with a user's position and consequently the user (Zhang et al. 2012).

Fig. 11. Medusa (Annett et al. 2011). The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. The plots show that Medusa cannot meet the minimum requirements of either scenario.

authors report that the system can robustly track two people along each long side of the table and one person along each short side. However, no systematic evaluation of the tracking accuracy has been reported other than the authors' description of cases where arms and touch points are incorrectly tagged.

Checkpoint: The Medusa plot (Figure 11) shows that Medusa does not meet the requirements of Scenario 1 (blue dashed plot) as it both relies on hardware installed in the environment (a frame around the table) and works only with large surfaces. It also does not meet the requirements of Scenario 2 (red dotted plot) as it restricts users' movement, position of interaction, and cannot uniquely identify the teacher.

For vision-based tabletop systems, Zhang et al. (2012) use the orientation of a touching finger, obtained from the touching hand's contour, to associate a touch with a user's position and consequently the user. The approach relies on the observation that the orientation of fingers resulting from distinct user positions around the tabletop has minimal overlap even when users are standing in proximity. In a study, where users were asked to use their index fingers only, the authors

report accuracies above 92% in challenging conditions (such as two-handed tasks) up to 98% in simple situation. While the approach is reported as simple, flexible and accurate, the need for a tabletop system that can detect hand contours to help calculate finger orientation (i.e., *embedded infrastructure* that is *depending on a specific sensing technology*) greatly limits its applicability. It also requires users to maintain their position around the table throughout the session (*restricted movement*), and only *differentiates rather than uniquely identifies users*. However, it should support *online, two-handed+, one-finger* interactions *anywhere* on the table and should work for *any surface* big enough for the hand shadow algorithm to work. Bootstrapper (Richter et al. 2012) uses a depth camera to recognize users by their shoes. While the main goal of the system is to recognize users working individually on a tabletop, the article briefly discusses using an approach similar to that of Zang et al. by locating hovering arms in the tabletop image then tracing them back to the edge of the table where they get associated with the recognized user located at that edge.

Checkpoint: Zhang et al.'s (2012) system (Figure 12) fails to meet the requirements of Scenario 1 (blue dashed plot) as it requires both a vision-based sensing system and a large surface. It also fails to meet the requirements of Scenario 2 (red dotted plot) as it restricts users' movement and cannot uniquely identify the teacher.

Group Touch (Evans et al. 2017) uses statistical modeling of touch pairs looking at the differences in orientation, distance, and time between two touches to determine whether they were performed by the same person or not. This means that the touch sensing technology used must allow for the calculation of touch orientation (vision-based Microsoft PixelSense tabletop was used in the studies). This approach results in a number of limitations such as only being able to know if one user's touches are different from another's if touches are happening simultaneously or within a very short period of time. They thus call their approach group touch as it groups sequences of touches belonging to the same user together. Accordingly, Group Touch supports a *limited user differentiation*, works best when around *large surfaces* (tabletops) where users are not expected to frequently jump around different areas of the screen thus imposing some *constraints on both user mobility and points of interaction*. While the hardware dependency, limited user differentiation support and constraints on mobility, positions of interaction, and surface size makes this approach not suitable for many practical scenarios including both educational scenarios used in this survey, Group Touch's most useful application is resolving conflicting inputs that occur during multi-user simultaneous interaction around large surfaces (i.e., attempting to drag the same object in two different directions, or closing a window/document while another user is interacting with its content). This approach has been categorized as zone-based due to its reliance on the spatial (as well as temporal) characteristics of touches to support differentiation.

4.4 Wearables and Handheld

For wearable-based user identification systems ((Vu et al. 2012; Roth et al. 2010; Meyer and Schmidt 2010; Marquardt et al. 2010; Marquardt et al. 2011)), it is possible to *uniquely identify* the wearable (and accordingly the user) *across devices*. However, in most cases, such identification usually depends on certain sensing technologies (capacitive (Vu et al. 2012), IR (Roth et al. 2010; Meyer and Schmidt 2010), detection of fiducial markers (Marquardt et al. 2010)). Yet, all of these approaches work with *any screen size*, allowing users to *move freely*, allowing *interaction anywhere* on the device, though some impose restrictions both on the position (Meyer and Schmidt 2010) and speed of interactions.

Specifically targeting *capacitive* touchscreens, Vu et al. (2012) use a small transmitter embedded into a ring for user identification (Figure 13(a)). The proposed approach does not require any hardware or firmware modification to off the shelf surfaces. When pressed against the capacitive

Fig. 12. Zang et al.'s (2012) system. The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. The qualifiers (values in brackets) show that the system cannot meet the requirements of Scenario 1 (the use of vision-based system and the need for a surface size that is larger than a tablet). The plots also show that that system cannot meet the minimum requirements of Scenario 2 as well.

(a)

(b)

(c)

Fig. 13. Wrist worn systems. (a) Vu et al.'s (2012) system with a prototype ring. (b) Meyer and Schmidt's (2010) advanced allocation using wristband orientation. (c) Fiduciary-tagged glove (Marquardt et al. 2010).

screen, the ring acts as a voltage source that creates a sequence of time-stamped touch events based on the ring's unique code. Due to practical limitations, a slow bit rate is used (only 4–5 bits per second) which means that short gestures are less likely to be accurately detected (detection rate falls below 85% for gestures shorter than 400 millisecond). To be able to detect the code transmitted by the ring, the *ring must be in direct contact with the screen*. However, it is possible to detect the presence of “a ring” if the *finger with the ring* touches the screen. Accordingly, in scenarios with only two users interacting with a device using one finger each (one with a ring and the other without), it is possible to identify touches resulting from the finger with the ring, thus making it possible to simply *differentiate between two users*.

Roth et al. (2010) also use a ring-like device but in this case it transmits a continuous pseudorandom IR pulse sequence. Their system thus *depends on hardware that can detect IR signals*. Each IR pulse sequence is associated with a particular user and all *touches in direct vicinity of the detected sequence* are associated with that user. The authors report on initial testing of the system with a data rate of 5 bits per second. The system took, on average 2.4 seconds to achieve authentication probability of 0.997 (8 data bits). The maximum was 4.8 seconds and the minimum was 1.8 seconds. With this *rather slow rate* and identification based on touches being near the detected sequence, the *system is designed with authentication in mind rather than identifying users with every touch in a typical multi-user interaction scenario*.

Meyer and Schmidt (2010) use Infra-Red (IR) pulsating wristbands (Figure 13(b)). It thus *depends on hardware that can detect IR signals*. Each wristband transmits a unique code, allowing a specific wristband to be associated with each touch registered in an area near the detected IR pattern. Since the wristband is placed at a distance from the touch points, two Light Emitting Diodes (LEDs) using a specific blinking pattern are utilized to determine a wristband's orientation, and thereby narrowing down the area of finger touch association based on the estimated position and orientation of the hand. This does impose some level of *constraints on the position of interaction* to ensure it is within the region associated with the wristband, but it also means that it allows *multi-touch interaction* within that region. No user study and performance measures are reported.

Marquardt et al. (2011, 2010) used gloves tagged with fiduciary markers (Figure 13(c)) to identify not only who touched the surface, but which part of the hand and which side of the hand that touched it (*all-fingers+*, and *two-hands+ if users wear two gloves*). Even though this approach requires users to wear special gloves and is touch sensing *technology dependent* (only works with hardware that can detect fiduciary marker), it is a reliable way of getting user input.

Checkpoint: Figure 14(a) shows the plots for the different wearable-based user identification systems. However, they all show dependence on a specific type of technology. As Scenario 1 specifically assumes capacitive touch sensing, only the system proposed by Vu et al. (2012) (dark gray) can meet the requirements. This, however, has some practical limitations in terms of instantaneity. For Scenario 2, the systems by Roth et al. (2010), Meyer and Schmidt (2010), and Marquardt et al. (2010) meet the requirements but those by Roth et al., and Myer and Schmidt may impose some restrictions due to their slower user identification performance.

Two techniques that are not specifically designed for user identification but can, in principal, support user identification are those proposed by Bianchi and Je (2017) and Webb et al. (2016). Both use wearables for user identification, but neither provide any type of study to evaluate the performance of their systems for user identification. Bianchi and Je (2017) use a wearable ring (Vib-Ring) that can augment touch information with an invisible identifier. This is achieved by modulating a structured vibration pattern that is transmitted through the finger to the touch device. This vibration pattern is then detected by the touch device's accelerometer. The vibration signal is then filtered and processed to extract the unique identifier. As the focus of this work is on touch

Fig. 14. (a) Wearable technologies vary in their support for user identification depending on the specifics of the approach used (Vu et al. 2012; Roth et al. 2010; Meyer and Schmidt 2010; Marquardt et al. 2010). (b) The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. Comparing the plots in (a) and (b) shows that while only the plot for Vu et al. (2012) (dark gray region) meets the capacitive touch sensing technology requirement, it may not meet the instantaneity requirements for all applications. On the other hand, the plot for Scenario 2 falls within the plots of Roth et al. (2010), Meyer and Schmidt (2010), and Marquardt et al. (2010) (depending on the instantaneity requirements).

augmentation rather than user identification, the system has a number of limitations if it is to be used for user identification. Apart from requiring a device that has an accelerometer (*hardware dependance*) which probably limits the use to small handheld devices (*size restriction (small)*), the system does not support simultaneous touches and suffers from a significant lag time (between off stage and vibration) with an average of 42 millisecond thereby limiting the usable frequency range to 4–20Hz and thus reducing its performance on the instantaneity axis. The system also requires the user to press and hold a button on the ring (to activate the vibration) while touching the surface.

Likewise, the system proposed by Webb et al. (2016) focuses on supporting bimanual interaction and requires a user to have a wearable motion sensor on each hand. In their study, a wrist band was worn on the non-dominant hand and a ring on the dominant hand. The system uses the accelerometer signal from the wearable devices to detect spikes (hard-contact forces) resulting from touching the display. If such a spike occurs within a specified time frame following a touch event, that touch event is associated with the hand wearing the corresponding sensor. In principle, this technique independent from the touch sensing technology, works with any screen size, can uniquely identify users across devices, and does not provide limitations on interaction position. It supports two-handed+ interaction as long as the user has a wearable sensor on each hand. However, the presented work focuses on the application for this technique for different bimanual interactions performed by one user and does not provide any study evaluating its use/suitability for user identification. It is not clear how the system performs for drag gestures where there may not be a clear “hard-contact” signal, when the users use soft taps. It is also not clear whether the ring performs better than the wrist band as the wrist band accelerometer may not necessary pick up the tap signal as clearly as a ring placed directly on that finger.

Fig. 15. (a) PhoneTouch enables direct phone interaction on surfaces in addition to conventional multi-touch (Schmidt et al. 2010a). (b) ShakeID requires a Microsoft Kinect camera in addition to the multi-touch display and phones with accelerometers (Rofouei et al. 2012).

Schmidt et al. (2010a) and Rofouei et al. (2012) used mobile phones with motion sensors to support some level of identification as opposed to wearables. PhoneTouch (Schmidt et al. 2010a) (Figure 15(a)) enables the use of phones to select targets on the surface by direct touch. Correlation in time between phone sensor (accelerometer) data and touch data from the surface is used to identify taps carried out with the phone. Phones and the interactive surface use Bluetooth to exchange synchronization messages and time stamps. To allow for user identification, *users must interact with the surface using their phones* (as a stylus) rather than their fingers. However, with phones being near ubiquitous personal devices, this approach allows for *unique user identification across devices* + . A vision-based tabletop is used to discriminate between phone taps and finger touches, so the reported implementation was *dependent on a specific type of hardware* that can both detect phone taps and distinguishes them from finger taps. This approach might be useful when infrequent identification is needed, but it is not a practical alternative during normal interaction with touch surfaces.

ShakeID (Rofouei et al. 2012) (Figure 15(b)) associates touches on vertical interactive displays to both individuals and their mobile devices using depth camera-based body tracking and real time device accelerometer data (i.e., the approach combines *environment*, and *worn/held device infrastructure*). A Microsoft Kinect (pointed at the users) is used to track users' hand positions which are then associated with touch contact positions (as with other vision-based system), the combination of which being sufficient for distinguishing between users. However, if the aim is to *uniquely identify users* + (or to link a touch to a device), the user needs to hold a motion sensor equipped mobile device that is paired to the system. ShakeID then correlates the acceleration for each paired device with the acceleration of all hands tracked by the Microsoft Kinect. The system can consequently associate each phone to a specific user's hand. As the system tracks all users' hands, the system can still associate a device to a user if the device is held with the hand that is not touching the screen if that hand is still making enough movements to allow for accurate correlation between the motion sensed by the device and that tracked by the camera (i.e., two-handed + with constraints). Due to practical limitations, the system requires users to *not stand too close to the display* (some *constraint on movement*) for the Kinect to acquire accurate data. Another limitation associated with the Kinect SDK is that the system works with two users only. In a small user study (7 pairs of participants) 94% and 92% accuracy was observed (depending on whether the users are working in parallel and side-by-side, or in close collaboration). The results, in addition to

Fig. 16. Using handheld devices. (a) PhoneTouch (Schmidt et al. 2010a). (b) ShakeID (Rofouei et al. 2012) in solid gray. The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2.

issues related to visual occlusion and field of view of the camera, depend on whether the users are holding the device in the hand that is touching the display or not, and in the level of movement of the hand holding the device.

While neither system meets the requirement for Scenario 1, both systems meet the requirements of Scenario 2 (despite some mobility limitations for ShakeID) as long as it is acceptable to hold a mobile device during the interaction.

Checkpoint: The visualizations in Figure 16(a) show that PhoneTouch is a touch sensing dependent approach, which utilizes computer vision. The need for cameras in addition to the need to interact with the surface using a phone makes this approach unsuitable for Scenario 1. ShakeID (Figure 16(b)), while achieving device independence, works only with large surfaces and requires hardware setup in the environment (body tracking cameras) which also makes it unsuitable for Scenario 1. As for Scenario 2, the plots show that, from a pure technical perspective, both PhoneTouch and ShakeID fully contain the plot of Scenario 2. PhoneTouch, however, requires that the interactive display on the wall uses vision-based sensing and that users are comfortable physically interacting with the surface using their phones, which is not very practical. Moreover, ShakeID requires that it is acceptable to hold a mobile device while interacting with the wall display.

4.5 Biometrics

Fiberio (Holz and Baudisch 2013) (Figure 17) is a special touch screen that authenticates users biometrically using fingerprints for each touch interaction (*embedded infrastructure*). The system was reported to support *online* finger print processing at an “interactive rate” (21 millisecond processing time per frame). In a small study that involved comparing 30 fingers (10 users), the cross-validated analysis resulted in 148 of 150 fingerprints being correctly matched, 0 wrong matches, and 2 no matches (for cases where users properly used their finger pads). For such a sample, the average processing time for matching a set against all others was 267 millisecond. However, to make this possible, the system relied on the use of a special large 4233dpi fiber optic plate with 40 million optical fibers as screen material, combined with rear-projection and a high-resolution camera thus requiring *custom sensing hardware*. However, the advantages are *unique user*

Fig. 17. Fiberio, a custom built touch screen that allows for authenticating users biometrically using fingerprints (Holz and Baudisch 2013).

identification across devices + with such support, *two handed*+, *multi-finger*+ *interaction* (if all fingers are registered), reasonable *online* speed, *freedom of movement*, and ability to *interact anywhere* on the surface. Practical considerations put *some constraints on screen size* as to how small or how large it can be.

Checkpoint: Figure 18 shows that the only axis where Fiberio cannot meet the requirements of Scenario 1 is touch sensing technology dependence. For Scenario 2, the only constraint for using Fiberio is the surface size and the practicality of building a large surface that can still achieve reliable fingerprint identification (the reported implementation is for a 19 in. display).

The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. Due to the need for custom hardware Fiberio is not suitable for Scenario 1. The plot for Scenario 2 shows that the only constraint on the use of Fiberio is display size.

Harrison et al. (2012) and Holz and Knaust (2015) relied on the impedance profile of users for biometric identification. The former approach, which is referred to as “Capacitive Fingerprinting” (Figure 19(a)), uses a sensing method based on Swept Frequency Capacitive Sensing (Sato et al. 2012), which measures the impedance of a user to the environment (ground in this case) across a range of AC frequencies. The technology supports real-time (*online*) user identification for *two-handed*, *multi-finger* interaction and shaped contacts such as palm press. Initial calibration of the “first touch” is required but it is reported to require less than 1 second. This approach also overcomes limitations associated with computer vision-based approaches such as occlusion and lighting and allows users to *move freely*, and *interact anywhere* on the surface. However, while the proposed technique does not require users to wear any devices, works with *any screen size*, and touch devices can easily be instrumented to support it (*embedded infrastructure*), the reported implementation is *technology dependent* requiring passive sensing technology (such as the use of IR driven touch panels) rather than the more widely used capacitive devices due to interference. Furthermore, it can only differentiate among *a small set of concurrent users and most importantly*

Fig. 18. Fiberio, fingerprint-based user identification (Holz and Baudisch 2013). The blue dashed plot represents the minimum requirements of Scenario 1 while the red dotted plot represents the minimum requirements for Scenario 2. Due to the need for custom hardware Fiberio is not suitable for Scenario 1. The plot for Scenario 2 shows that the only constraint on the use of Fiberio is display size.

Fig. 19. (a) Capacitive Fingerprinting system (Harrison et al. 2012). (b) Biometric Touch Sensing (Holz and Knaust 2015).

Fig. 20. (a) Capacitive Fingerprinting (Harrison et al. 2012). (b) Biometric Touch Sensing in hash pattern (Holz and Knaust 2015). The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. The plot for Scenario 1 falls within that of Holz and Knaust, but not that of Harrison et al. due to the specific sensing hardware restrictions for that system. Likewise, the plot of Scenario 2 falls within that of Holz and Knaust but not that of Harrison et al., as the latter cannot uniquely identify the teacher.

users can only touch the screen sequentially and not simultaneously. The authors also reported a limitation with persistence identification. This is because variability in an individual can, sometimes, be larger than that between individuals. Accordingly, this approach is only recommended for scenarios with ad-hoc interaction for relatively short periods where *user differentiation* (rather than unique identification) is useful but not critical.

Checkpoint: The plot for the Capacitive Fingerprinting (Figure 20(a)) shows that this approach satisfies all the requirements of Scenario 1 apart from the touch sensing technology. However, since commonly used tablets do not meet the “passive-sensing” requirement due to interference, the system is not suitable for Scenario 1. Similarly, the plot shows that the approach also does not meet the identification of the teacher as a unique individual; thus, it does not meet the requirement of Scenario 2.

Biometric + Wearable: To overcome many of the limitations in Harrison et al.’s (2012) approach, and specifically targeting the more widely used capacitive screen (*technology dependent*) Holz and Knaust (2015) proposed “biometric touch sensing.” This technique uses a custom-built *wrist-worn* device “Bioamp” that senses the impedance profile of the user’s wrist and modulates a code onto the user’s body through skin using periodic electrical signals. The capacity measured upon touch is affected by this signal thus allowing the transmitted code to be read. Due to the low update rate of current commodity devices, it is not possible to send the actual biometric information onto the user’s body to the device. The alternative is to send this information, along with a unique identifier code to the device using Bluetooth which also gets encoded and sent to the device worn by the user. The device then looks up the biometric information to link it with one of the pre-registered users and the associated touch input event. This approach does not require any hardware changes but does require changes at the operating system level. The focus of this work was on user authentication upon each touch and eliminating the need for users’ logins and use of passwords, rather than on user identification for simultaneous multi-user scenarios. However, the identification is

reported to be “near instant” (*online*) and the technique does support multiple simultaneous users, which was demonstrated from the perspective of authentication—where more than one person is required to confirm certain activities (*unique identification across supported devices*). It also supports *multi-finger* interaction, allows users to *move freely*, *interact anywhere* on the surface and works with capacitive screens of *any size*. One limitation reported with regards to multi-user support is that no more than one user can touch the same sensing line. However, in practice, the pitch between sense lines is normally small enough such that the likelihood of this happening is very small. Accordingly, the main shortcomings of this approach seem to be the need for a custom built wearable device, working only with capacitive touch screen, and the need for operative system level changes to interactive surfaces which may or may not be problematic depending on the context of use.

Checkpoint: Figure 20(b) shows that the biometric touch sensing approach meets all the requirements of Scenario 1 (blue dashed plot). It only requires that students use wrist worn devices to support identification (which the plot shows to be an acceptable compromise). What the chart does not show is that it also requires changes at the operating system level, which is not always possible in a school setting. For Scenario 2 (red dotted plot), the plots also show that if the interactive display used in the classroom is capacitive, then this approach meets all the requirements.

Biometric + Zone-based: HandsDown (Schmidt et al. 2010b) and MTi (Blažica et al. 2013) can uniquely identify users’ hands when placed flat on a surface. They are not proposed as general approaches for directly identifying and associating touches to users but by identifying a user’s hand, it is possible to define territories associated with that user in a customizable region around the identified hand. Accordingly, they allow *one-handed*, *multi-finger* interaction with *constraints on interaction position*. These approaches also require a *large surface* to allow for space for a flat hand as well as a region around it for interaction. Both approaches rely on biometric user authentication but, as the authors suggest, not for security critical contexts and they only work well for a small number (10–100) of pre-registered users (*limited unique identification across devices*).

HandsDown (Schmidt et al. 2010b) can identify users who put their hands down flat on any interactive surface that can detect arbitrarily shaped objects in addition to finger touches (mostly camera based systems, *embedded infrastructure* and *sensing hardware dependent*). Once the user’s hand is placed flat on the surface, characteristic features of the hand contour are extracted from the images captured by the surface and then classified using Support Vector Machines (SVM). This means that for users to be identified, they need to first register with the system by providing sample hand contours. The authors reported that for scenarios of 10 known and 0 unknown users, an acceptance rate of 92% and false acceptance rate of 0.1% can be achieved. A different threshold would achieve 97% acceptance rate but with 1% false acceptance rate. The performance drops for scenarios where unknown users are expected to use the system.

Still relying on hand biometric and overcoming some of the limitations of HandsDown, Blažica et al. (2013) proposed MTi, a method for user identification based on features obtained only from the coordinates of the five touch points of a user’s hand. Accordingly, this approach can be used with any type of touch technology (*sensing technology independent*) *large enough* to accommodate a hand and that can detect at least five touches simultaneously. Like in HandsDown, users are required to place their hands on the table, but with fingers comfortably stretched apart, to be uniquely identified. However, this approach is susceptible to the geometry of the whole hand, the quality of both training and test data, and the size of the training set. For a data set of 34 users, the authors reported 94.69% identification accuracy and for 100 users, MTi achieved 94.33%

Fig. 21. MTi (light gray plot) (Blažica et al. 2013) works on any surface that can detect more than five touch points. HandsDown (dark gray plot) (Schmidt et al. 2010b) requires hardware that can detect arbitrarily shaped objects. The blue dashed plot represents the minimum requirements of Scenario 1, while the red dotted plot represents the minimum requirements for Scenario 2. Due to the need for a space larger than a flat hand, neither system is suitable for Scenario 1. The plots also show that due to the limitation on interaction position neither system is suitable for Scenario 2 as well.

accuracy. As reported by the authors, unless combined with other approaches, MTi cannot be used for continuous tracking of the touch points.

Checkpoint: HandsDown and MTi have very similar plots (Figure 21). But since both require a screen that is larger than a flat hand, neither of them is suitable for pairs working on tablets as in Scenario 1. For Scenario 2, the weakness identified from the plots is the restriction on interaction location. Unless a “significant” compromise is made such that identification is only required for the teacher for specific commands, neither technique is suitable for Scenario 2.

5 DISCUSSION

An ideal user identification system should not require any external hardware infrastructure to work, should function independently of the sensing technology, can uniquely, and reliably, identify

Fig. 22. Existing systems make compromises to achieve identification for a specific hardware or use cases.²

users across devices, support handedness and even identify the finger interacting with the surface, works instantly, should not restrict users' movement nor positions of interaction, and should work with any screen size.

The visualizations made possible by the proposed unified model (Figure 22), make it clear that none of the existing approaches satisfy all these "ideal" requirements.

Compromises are made based on the specific problem to address or the context to work in. One of the main objectives of the model is to help practitioners select a system that best fits their requirements. The visualizations, as demonstrated in the various checkpoints for the two proposed scenarios, can be a good starting point for identifying which solutions to look at more closely and which are not well suited to the application scenario at hand.

²Visualizations of all the systems covered in this survey can be downloaded as SVG vector graphics from <https://github.com/ankharrufa/umodel-tuid.git>.

For example, achieving sensing technology independence often involves having to use external hardware in the environment with large interactive surfaces that are normally fixed in place (Dohse et al. 2008; Clayphan et al. 2013; Murugappan et al. 2012; Barab et al. 2004; Annett et al. 2011). Accordingly, this is not always a practical option. The approaches that achieved sensing technology independence without external hardware are zone-based. These make a significant compromise by requiring all interactions to be within pre-defined regions associated with each user (Scott et al. 2004), or at best dynamically controlled regions (Blažica et al. 2013). These also generally have limitations with regards to handedness and uniquely identifying users.

The practical alternative has been to propose solutions that, while dependent on technology, work on widely used ones such as capacitive surfaces (Vu et al. 2012; Holz and Knaust 2015) rather than less common vision-based ones for example. To achieve this, however, some sort of a wearable device is normally involved. This can be a ring that users need to use as a medium of interaction (Vu et al. 2012), or a more practical alternative where users interact freely with the capacitive screen while wearing a custom built wrist band (Holz and Knaust 2015). A promising research direction is looking at the use of motion sensing wristbands to support user identification by correlating touch data with accelerometer data. The increasing popularity of accelerometer equipped smart watches and fitness trackers has already lead to these devices being used to enrich (or increase the expressiveness of) touch interactions, with touch surfaces independent of the sensing hardware used (Wilkinson et al. 2016).

Also from the visualizations, we can quickly see that to achieve the goal of uniquely identifying users across devices, the compromises are either relying on a custom built hardware for supporting fingerprint identification (Holz and Baudisch 2013) (i.e., only cross devices that use this custom hardware); use custom built worn devices that only work with some, but not all, sensing technologies (Holz and Knaust 2015; Roth et al. 2010; Meyer and Schmidt 2010; Marquardt et al. 2010); or use handheld devices (such as phones which are used as the unique identifier). For handheld devices, users have to interact indirectly through the handheld device (Schmidt et al. 2010a) or require extra hardware in the environment (cameras) to still support direct interaction (Rofouei et al. 2012).

Similar compromises are evident from the graph when it comes to the other axes. Accordingly, what the model and its visual representation can help with is to make it easy to specify the minimum requirements of a system on the graph (e.g., Figure 4). Any user identification technique whose plot fully contains the requirement plot (taking the qualifiers into account) is considered adequate.

By means of the different checkpoints comparing existing systems to two possible classroom scenarios, it was demonstrated that no system directly satisfies all the requirements of either scenario without making some compromises. For example, since the requirement to work on off-the-shelf capacitive devices with the ability to interact anywhere on the relatively small tablet screens are essential, this quickly discards all the reviewed systems with embedded infrastructure as they either require different type of hardware or are zone-based. With overhead cameras being impractical in schools, the compromise to be made then is to use worn devices to support identification. With this compromise, the only system that fits the requirements is that of Holz and Knaust (2015). On the other hand, Scenario 2 is more forgiving in terms of allowing for custom built hardware and have infrastructure in the environment. However, it has the essential requirements of supporting a large display size, interaction anywhere on the display, user mobility and uniquely identifying the teacher. These immediately exclude many of the reviewed systems simply by looking at their plots. The remaining ones are largely camera based ones, wearables, and again Holz and Knaust (2015), and the decision among these will depend on the display hardware to be used.

Another significant benefit of the unified model is to researchers working on new user identification techniques on multi-touch surfaces. The model's terminology and radar-chart visualization provide researchers with a systematic method for reporting their work making it easy for readers to position this work within existing systems. In addition to its reporting power, the model also helps researcher in evaluating their work at both the early stages during idea development (a form of discount evaluation (Pinelle and Gutwin 2002)) and the final stages by ensuring that the researchers evaluate and report on all the parameters highlighted by the unified model.

A limitation in the visualization relate to surface size and hardware independence. For surface size, while moving towards the center reflects a restriction on surface size, we need to specify whether the restrictions are for supporting smaller surfaces only or larger ones. Also for touch sensing dependence axes, and specifically for the "dependent" point on the axes, we must textually express the type of hardware the technique is dependent on. For both these cases, the text-based qualifiers must be treated as an integral part of the visualization. The only way around that is to increase the number of axes which will significantly complicate the plots, rendering them impractical to use.

The survey conducted in this article highlights the problem-specific, or hardware specific, nature of most of the reviewed system. Many of the techniques are reported to have achieved very high identification accuracy and thus are more than suitable for the specific problem addressed. Using a technique that does not match well the application at hand, however, can significantly affect the performance level and that is where visualizations like the ones proposed by this unified model can be invaluable. While it is difficult to anticipate the direction of development of touch sensing technology, it is reasonable to expect that with faster processors (and frame rate for vision based systems) the instantaneity performance of some of the systems discussed here will improve. Higher sensing resolutions (or higher camera resolutions for vision based systems) have the potential to improve user differentiation (moving from differentiation to identification or providing more reliable identification) and the supported points of interaction (ability to distinguish between users' fingers and hands). On the other hand, parameters such as touch sensing dependence, infrastructure, surface size, interaction position, and (to a lesser extent) user mobility are only dependent on the specific characteristics of the approach used and are not affected by any possible improvements in hardware capabilities.

REFERENCES

- A. Al-Qaraghuli, H. B. Zaman, P. Olivier, A. Kharrufa, and A. Ahmad. 2011. Analysing tabletop based computer supported collaborative learning data through visualization. In *Visual Informatics: Sustaining Research and Innovations. IVIC 2011*, H. Badioze Zaman et al. (Eds.). Lecture Notes in Computer Science, vol. 7066. Springer, Berlin, Heidelberg.
- Michelle Annett, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2011. Medusa: A proximity-aware multi-touch tabletop. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST'11)*. ACM, New York, NY, 337. DOI: <https://doi.org/10.1145/2047196.2047240>
- Sasha Barab, Steve Schatz, and Rebecca Scheckler. 2004. Using activity theory to conceptualize online community and using online community to conceptualize activity theory. *Mind Cult. Act.* 11, 1 (February 2004), 25–47. DOI: https://doi.org/10.1207/s15327884mca1101_3
- BESA. 2013. *Tablets and Apps in Schools 2013*. BESA.
- Achintya K. Bhowmik. 2014. *Interactive Displays: Natural Human-Interface Technologies*, 1st ed. Wiley Publishing.
- Andrea Bianchi and Seungwoo Je. 2017. Disambiguating touch with a smart-ring. In *Proceedings of the Augmented Human International Conference*. ACM, 27.
- Bojan Blažica, Daniel Vladušić, and Dunja Mladenčić. 2013. MTi: A method for user identification for multitouch displays. *Int. J. Hum. Comput. Stud.* 71, 6 (2013), 691–702. DOI: <https://doi.org/10.1016/j.ijhcs.2013.03.002>
- Andrew Clayphan, Roberto Martinez-Maldonado, Christopher Ackad, and Judy Kay. 2013. An approach for designing and evaluating a plug-in vision-based tabletop touch identification system. In *Proceedings of the 25th Australian Computer-Human Interaction Conference on Augmentation, Application, Innovation, Collaboration (OzCHI'13)*. ACM, New York, NY, 373–382. DOI: <https://doi.org/10.1145/2541016.2541019>

- Paul Dietz and Darren Leigh. 2001. DiamondTouch: A multi-user touch technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology (UIST'01)*. ACM, New York, NY, 219. DOI : <https://doi.org/10.1145/502348.502389>
- K. C. Dohse, Thomas Dohse, Jeremiah D. Still, and Derrick J. Parkhurst. 2008. Enhancing multi-user interaction with multi-touch tabletop displays using hand tracking. In *Proceedings of the 1st International Conference on Advances in Computer-Human Interaction*. IEEE, 297–302. DOI : <https://doi.org/10.1109/ACHI.2008.11>
- Alan Esenther, Cliff Forlines, Kathy Ryall, and Sam Shipman. 2002. Diamondtouch sdk: Support for multi-user, multi-touch applications. *Mitsubishi Electric Research Laboratories, Inc.*, (2002).
- Abigail C. Evans, Katie Davis, James Fogarty, and Jacob O. Wobbrock. 2017. Group Touch: Distinguishing tabletop users in group settings via statistical modeling of touch pairs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI'17)*. ACM, New York, NY, 35–47. DOI : <https://doi.org/10.1145/3025453.3025793>
- Chris Harrison, Munehiko Sato, and Ivan Poupyrev. 2012. Capacitive fingerprinting: Exploring user differentiation by sensing electrical properties of the human body. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST'12)*. ACM, New York, NY, 537. DOI : <https://doi.org/10.1145/2380116.2380183>
- I. Heemskerk, E. Kuiper, and J. Meijer. 2014. Interactive whiteboard and virtual learning environment combined: Effects on mathematics education. *J. Comput. Assist. Learn.* 30, 5 (October 2014), 465–478. DOI : <https://doi.org/10.1111/jcal.12060>
- Christian Holz and Patrick Baudisch. 2013. Fiberio: A touchscreen that senses fingerprints. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST'13)*. ACM, New York, NY, 41–50. DOI : <https://doi.org/10.1145/2501988.2502021>
- Christian Holz and Marius Knaust. 2015. Biometric touch sensing: Seamlessly augmenting each touch with continuous authentication. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST'15)*. ACM, New York, NY, 303–312. DOI : <https://doi.org/10.1145/2807442.2807458>
- Ahmed Kharrufa, James Nicholson, Paul Dunphy, Steve Hodges, Pam Briggs, and Patrick Olivier. 2015. Using IMUs to identify supervisors on touch devices. In *Human-Computer Interaction*. Springer, Cham, 565–583.
- Ahmed Kharrufa, Madeline Balaam, Phil Heslop, David Leat, Paul Dolan, and Patrick Olivier. 2013. Tables in the wild: Lessons learned from a large-scale multi-tabletop deployment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*. ACM, New York, NY, 1021. DOI : <https://doi.org/10.1145/2470654.2466130>
- Ahmed Kharrufa, David Leat, and Patrick Olivier. 2010. Digital mysteries: Designing for learning at the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'10)*. ACM, 197.
- Ahmed Kharrufa, Roberto Martinez-Maldonado, Judy Kay, and Patrick Olivier. 2013. Extending tabletop application design to the classroom. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces (ITS'13)*. ACM, New York, NY, 115–124. DOI : <https://doi.org/10.1145/2512349.2512816>
- Nicolai Marquardt, Johannes Kiemer, and Saul Greenberg. 2010. What caused that touch?: Expressive interaction with a surface through fiduciary-tagged gloves. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'10)*. ACM, New York, NY, 139. DOI : <https://doi.org/10.1145/1936652.1936680>
- Nicolai Marquardt, Johannes Kiemer, David Ledo, Sebastian Boring, and Saul Greenberg. 2011. Designing user-, hand-, and handpart-aware tabletop interactions with the TouchID toolkit. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'11)*. ACM, New York, NY, 21. DOI : <https://doi.org/10.1145/2076354.2076358>
- R. Martinez, K. Yacef, J. Kay, A. Al-Qaraghuli, and A. Kharrufa. 2011. Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. In *Proceedings of the 4th International Conference on Educational Data Mining (EDM'2011)*. 111–120.
- Roberto Martinez, Anthony Collins, Judy Kay, and Kalina Yacef. 2011. Who did what? Who said that? In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'11)*. ACM, New York, NY, 172. DOI : <https://doi.org/10.1145/2076354.2076387>
- Tobias Meyer and Dominik Schmidt. 2010. IdWristbands: IR-based user identification on multi-touch surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'10)*. ACM, New York, NY, 277. DOI : <https://doi.org/10.1145/1936652.1936714>
- M. R. Morris, A. Cassanego, A. Paepcke, T. Winograd, A. M. Piper, and A. Huang. 2006a. Mediating group dynamics through tabletop interface design. *IEEE Comput. Graph. Appl.* 26, 5 (September 2006), 65–73. DOI : <https://doi.org/10.1109/MCG.2006.114>
- Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. 2006b. Cooperative gestures: Multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*. ACM, New York, NY, 1201. DOI : <https://doi.org/10.1145/1124772.1124952>
- Meredith Ringel Morris, Kathy Ryall, Chia Shen, Clifton Forlines, and Frederic Vernier. 2004. Beyond “social protocols”: multi-user coordination policies for co-located groupware. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative work (CSCW'04)*. ACM, New York, NY, 262. DOI : <https://doi.org/10.1145/1031607.1031648>

- Sundar Murugappan, Vinayak, and Karthik Ramani. 2012. Extended multitouch: Recovering touch posture and differentiating users using a depth camera. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST'12)*. ACM, New York, NY, 487. DOI: <https://doi.org/10.1145/2380116.2380177>
- David Pinelle and Carl Gutwin. 2002. Groupware walkthrough: Adding context to groupware usability evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Changing Our World, Changing Ourselves (CHI'02)*. ACM, New York, NY, 455. DOI: <https://doi.org/10.1145/503376.503458>
- Anne Marie Piper, Eileen O'Brien, Meredith Ringel Morris, and Terry Winograd. 2006. SIDES: A cooperative tabletop computer game for social skills development. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW'06)*. ACM, New York, NY, 1. DOI: <https://doi.org/10.1145/1180875.1180877>
- Raf Ramakers, Davy Vanacken, Kris Luyten, Karin Coninx, and Johannes Schöning. 2012. Carpus: A non-intrusive user identification technique for interactive surfaces. In *Proceedings of the 25th Annual ACM Symposium on User interface Software and Technology (UIST'12)*. ACM, New York, NY, 35. DOI: <https://doi.org/10.1145/2380116.2380123>
- Jaffar Razmi, Mohamed Zairi, and Yasar F. Jarrar. 2000. The application of graphical techniques in evaluating benchmarking partners. *Benchmarking An Int. J.* 7, 4 (October 2000), 304–314. DOI: <https://doi.org/10.1108/14635770010378936>
- Stephan Richter, Christian Holz, and Patrick Baudisch. 2012. Bootstrapper. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI'12)*. ACM, New York, NY, 1249. DOI: <https://doi.org/10.1145/2207676.2208577>
- Jochen Rick and Jochen. 2012. Proportion: A tablet app for collaborative learning. In *Proceedings of the 11th International Conference on Interaction Design and Children (IDC'12)*. ACM, New York, NY, 316. DOI: <https://doi.org/10.1145/2307096.2307155>
- Mahsan Rofouei, Andrew Wilson, A. J. Brush, and Stewart Tansley. 2012. Your phone or mine?: Fusing body, touch and device sensing for multi-user device-display interaction. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI'12)*. ACM, New York, NY, 1915. DOI: <https://doi.org/10.1145/2207676.2208332>
- Volker Roth, Philipp Schmidt, and Benjamin Gildenring. 2010. The IR ring: Authenticating users' touches on a multi-touch display. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology (UIST'10)*. ACM, New York, NY, 259. DOI: <https://doi.org/10.1145/1866029.1866071>
- K. Ryall et al. 2006. Identity-differentiating widgets for multiuser interactive surfaces. *IEEE Comput. Graph. Appl.* 26, 5 (September 2006), 56–64. DOI: <https://doi.org/10.1109/MCG.2006.108>
- Kathy Ryall et al. 2005. iDwidgets: Parameterizing widgets by user identity. In *Proceedings of the Human-Computer Interaction-INTERACT*. 1124–1128.
- Sidharth Sahdev et al. 2017. GhostID. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI'17)*. ACM, New York, NY, 15–27. DOI: <https://doi.org/10.1145/3025453.3025719>
- Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: Enhancing touch interaction on humans, screens, liquids, and everyday objects. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI'12)*. ACM, New York, NY, 483. DOI: <https://doi.org/10.1145/2207676.2207743>
- Dominik Schmidt, Fadi Chehimi, Enrico Rukzio, and Hans Gellersen. 2010a. PhoneTouch: A technique for direct phone interaction on surfaces. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology (UIST'10)*. ACM, New York, NY, 13. DOI: <https://doi.org/10.1145/1866029.1866034>
- Dominik Schmidt, Ming Ki Chong, and Hans Gellersen. 2010b. HandsDown: Hand-contour-based user identification for interactive surfaces. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries (NordicCHI'10)*. ACM, New York, NY, 432. DOI: <https://doi.org/10.1145/1868914.1868964>
- Dominik Schmidt, Ming Ki Chong, and Hans Gellersen. 2010c. IdLenses: Dynamic personal areas on shared surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'10)*. ACM, New York, NY, 131. DOI: <https://doi.org/10.1145/1936652.1936678>
- Stacey D. Scott, M. Sheelagh, T. Carpendale, and Kori M. Inkpen. 2004. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW'04)*. ACM, New York, NY, 294. DOI: <https://doi.org/10.1145/1031607.1031655>
- Reflective Thinking. Digital Mysteries for iPads. <http://www.reflectivethinking.com/digitalmysteries-ipad>
- Tam Vu et al. 2012. Distinguishing users with capacitive touch communication. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (Mobicom'12)*. ACM, New York, NY, 197. DOI: <https://doi.org/10.1145/2348543.2348569>
- Benjamin Walther-Franks, Jens Teichert, Markus Krause, and Marc Herrlich. 2008. User detection for a multi-touch table via proximity sensors. In *Proceedings of the IEEE Tabletops and Interactive Surfaces*.
- Andrew M. Webb, Michel Pahun, Ken Hinckley, and Bill Buxton. 2016. Wearables as context for guarded-biding bimanual touch. In *Proceedings of the Symposium on User Interface Software and Technology (UIST'16)*. ACM, 287–300. DOI: <https://doi.org/10.1145/2984511.2984564>

- Gerard Wilkinson et al. 2016. Expressy: Using a wrist-worn inertial measurement unit to add expressiveness to touch-based interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*. ACM, New York, NY, 2832–2844. DOI: <https://doi.org/10.1145/2858036.2858223>
- Hong Zhang, Xing-Dong Yang, Barrett Ens, Hai-Ning Liang, Pierre Boulanger, and Pourang Irani. 2012. See me, see you: A lightweight method for discriminating user touches on tabletop displays. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI'12)*. ACM, New York, NY, 2327. DOI: <https://doi.org/10.1145/2207676.2208392>

Received May 2017; revised September 2017; accepted September 2017